Comparision of three different approaches for forecasting accommodation estimations*

Mehmet Emin Akkaya
North Cyprus Near East University

Ihsan Tolga MEDENI
Yildirim Beyazit University

Abstract

In the tourism sector, estimation in accommodation numbers is crucial for the preparation of the enterprises to answer the demand of the customers. In this study, by looking at Turkey's three of the provinces, based on the highest tourist overnight stays, Linear Regression, MLP and SVR analyses methods has been implemented by looking at the annual data range of domestic and foreign overnight stay data. This is a comparative study for the provinces of Antalya, Istanbul, Mugla which has the highest tourist overnight stay numbers of Turkey by time series analyses performed by Linear Regression, MLP and SVR analyses methods. Annual data range has been used to estimate the tourism demand. Compared to the domestic and foreign overnight stay number estimation studies by 3 different regression methods. Used multivariate data and different tourism destinations. In the graphs with high density of cyclical fluctuation, it was observed that SVR method gave the closest result to real values. In the graphs with less density from high of cyclical fluctuation, it was observed that MLP method gave the closest result to real values. In the graphs with low density of cyclical fluctuation, it was observed that Linear method gave the closest result to real values. It is predicted that the results obtained from this study shall be useful for tourism personnel, researches, investors, tourism executives and tourism planning institutions that applies data mining techniques for tourism demand estimation applications.

Keywords : Accommodation Estimates, Tourism Forecasting, Time Series Regression

---

doi: https://doi.org/10.1043/48.2019.5/5.05

Introduction

Demand forecasting in tourism, provides clear vision and planning for the future. It gives you the opportunity to recognize changing market trends and keep up with the competition. It enables understanding of changes in the sector. Correct demand forecasts, meet incoming demand and reaching income targets are ensured. Savings from expenses. It prevents transfer of money to unnecessary resources. Strengthens liquidity. Allows the planning of discounts. It provides planning of the number of staff and cost. Estimating the periods when the demand will decrease and increase will allow us to periodically plan financing.

In the tourism sector, estimation in accommodation numbers is crucial. If there is no demand forecast in tourism, a clear vision and planning for the future can not be made. It does not give you the opportunity to recognize the changing market trends and keep up with the competition. The changes in the sector are incomprehensible. Incoming demand is unknown and income targets are not reached. Expenses can not be saved. It does not prevent money transfer to unnecessary resources. Liquidity weakens. Discounts do not allow planning. The number of staff and cost cannot be planned. Demand declines and increases can not be estimated periodically. Financing can not be planned periodically.
For the solution of the problem, the data belonging to the variables that will affect the demand forecasts in the tourism sector were collected and the demand estimation was done by the regression analysis method.

Data mining is the science of deducing useful data from huge amount of data. Estimations can be made on data types and data trends by data mining methods. This is especially useful to make conscious decisions on future of the enterprises. Data mining is used to estimate customer demand trends and increase sale performance by existing data on sale details. (Hemambika 2005) Correlation analysis is a statistical method used to test the linear relationship between two variables or relationship of one variable with two or more variables and measure level of this relationship if any. Data mining is generally used for statistical methods.(Kadambi 2005) Models used for data mining are examined in two main titles, estimative and descriptive. (Lim 2013) Data mining techniques have a wide range by options coming from many disciplines. These selections include techniques such as support vector machines, correlation, linear regression, non-linear regression, genetic algorithms, artificial neural networks, decision trees et cetera. Although size of the database is significant for natural solution of the problem, selecting a data mining technique is essential.(Adriaans etc. 1996: 57) In some cases, data should have higher scales. Following a more complex algorithm running efficiently on a data set of which dimension is reduced, a combined procedure a la

scalable algorithms can be predicted.(Cloyd 2002) Time series analysis is one of the most popular approaches used for estimates. Past data and trends are taken into consideration and used for estimating the future. (Al-Alwani 2014)

Time Series are digital quantities in which values of variables can be observed successively. It is not only a source of information used for estimating the future but also a method serving the same purpose. Time series consist of trends, seasonal fluctuations, cyclical fluctuations and irregular movements (error term). (Ibrahim 1995) Time series data are one of best recognized types of data. Generally, data are formed instantly and measured by real-time global measurements. (Rakthanmanon 2012) Time series estimates are equally spaced data which estimates by historical values formed in time and express these historical values by certain time intervals such as monthly sale data and daily electricity consumptions. (Seliem 2006) There are two different time series analyses which are the most popular ones: Modelling time series: time series forming mechanism to obtain insight and prediction time series: to estimate future values of the time series variable. (Yunyue 2004)

Regression methods are probably the oldest and most popular approach for exploring the problem of functional dependency of data (Draper etc. 1998), (Maimon etc. 2000). Diversity of regression methods, including the simple one, is related to the number of linear, multilinear and non-linear models. Multiple regression analysis is one of the most popular regression methods. This method includes studies related to linear relationship between univariate and multivariate clusters (Bello 1995).

Theoretically, increase in the quantity of data ensuring predictive regression estimate also increases score quality of equation estimation. It ensures better estimations. On other hand, missing values create inadequate and weak predictive estimations. (Frane 1976), (Sutarso 1995).

It shall be better to estimate parameters obtained from more observations for regression analysis. In case of a perfect match, estimated value is on par with the real value. However, if this is not the case, a difference between a certain real value and estimated value, in other words deviation shall be observed.(Draper etc. 1998), (Freund 1998), (Siripitayananon 2002) Data mining shall be used for statistical predictive estimates and studies such as price, sale, cost and demand in hotel and accommodation facility applications. (Ding 2014) In the studies on sale estimates, time series data analyses have been used and positive reflections have been observed in development of monthly sale method for sale and marketing. (Gaojun etc. 2009)

Regression models based on sample data sets have been emphasised in previous approaches for sale estimations. However, overly unrealistic results have arisen in this regression model. In most recent theoretical studies on statistics, support vector regression (SVR) method has been offered as a new method to overcome the problem of unrealistic (fabricated)results. In

contrast with the traditional regression model, the aim of SRV is to obtain empiric risk with minimum structural risk. In conclusion, it has been seen in the study that SVR is a superior method. (Yu etc. 2013) Support vector machine for data regression (SVR) has become a more powerful learning model by better generalization ability. (Kong etc. 2015) Based on statistical learning theory, SVR has been used for an efficient neural network algorithm and successful sale estimation methods to solve the problem of non-linear regression estimates. (Dai etc. 2015) Internal factors (seasonality, trend, promotions, price variances etc.) and external factors (cross sale, cannibalization effect, promotions of the competitor etc.) have been taken into consideration in the sale estimation studies performed by time series. (García etc. 2012) The results of sale estimates have generally been used to support purchasing and production decision. Long term stability is a significant concern for making a decision on purchasing and production. However, the sale departments generally want to react and update the estimation model more frequently in more efficient markets. There are two significant issues, accuracy of estimation model and currency of the estimation model, while making sale decisions related to the entire market. Long term stability and accuracy/currency are always significant for different decision types. Therefore, these two issues should be dealt in a more comprehensive research. (Chern etc. 2015) Better estimation models have been developed day by day. A new demand estimation system based on estimation model has been developed in more recent studies (to help wholesalers and retailers for issues like logistics distribution). Such data mining studies are significant to improve customer satisfaction (Fang etc. 2011) Performances of the estimation model have been compared by using the machine learning techniques known as ANN (Artificial Neural Network), SVM (Support Vector Machine) and LM (Linear Modelling). (Moon etc. 2014) Researches have been carried out on support vector machines in the discipline of medicine. (Wu etc. 2016) Actual estimative studies have been made by using the support vector regression model, most developed estimation method for data mining. (Tezel etc. 2016) Support vector regression has been firstly developed to solve the problem of classification within the frame of statistical learning theory (Burges 1998), (Cristianini etc. 2000). It has been used in estimation and predictive estimation researches due to its terrific estimation qualities. (DiPillo etc. 2016) Support vector regression is a useful method for creating an efficient sale estimation structure. (Dai etc. 2014) Support vector regression is an artificial intelligence estimation tool based on statistical learning theory and structural risk minimization principle. (Lu 2014)

Support vector regression (SVR) is also used for non-linear regression solution in time series problems and estimations. (Zhang etc. 2016). Support vector regression is a significant tool for time series estimations. Researches have been made on accuracy and reliability of estimations and SVR has been tested in terms of accuracy and reliability. (Zhao etc. 2013) Support vector machines have yielded positive results for studies aiming to increase analysis

verification rate. It has drawn the attention of the scientific community by these results. (Shawe etc. 2000), (Gomez etc. 2011) It has been seen that time series analyses and support vector regression method have been used in researches on sale estimations and estimation methods for data mining and, SVR (support vector regression) has had a better performance in multivariate time series analyses. (Khalil Zadeh etc. 2014), (Cankurt etc. 2016)

Tourism has provided a continuous growth and diversity in the last 60 years despite some obstacles having significant effects on tourist movements in the short term such as wars, regional epidemic and financial crises. Accurate demand estimates form the basis on which decisions on tourism and hotels are made regarding pricing and corporate strategies. At the same time, medium and long term tourism and hotel demand estimates are required for investment decisions of the private actors and public infrastructure investments. Demand modelling and estimation is a significant area for tourism and accommodation researches. (Wu etc. 2016).

Arrival of tourists at a destination is the traditional and most commonly used indicator of the tourism demand. Other two popular indicators are tourism expenditures (Cortés-Jiménez etc. 2011), (Smeral 2010) and number of overnight stays (Athanasopoulos etc. 2008), (Baggio etc. 2016)

Demand for accommodation at hotels is measured by different variables from various perspectives. Some variables such as arrivals of visitors (Guizzardi etc.2015) number of overnight stays (Falk 2014), (Lim etc. 2009) number of rooms sold and occupancy rate (Koupriouchina etc. 2014), (Wu etc. 2010) are related to the demand scale. (Song etc. 2011) In terms of data set, annual data have been used by many researchers for tourism and hotel demand modelling and estimation studies. These studies normally focus on factors affected by tourism (or hotel) demand and/or long term relationships between medium and long term trend estimates. (Guizzardi etc. 2015), (Song etc. 2011)
Artificial intelligent based models and artificial intelligent techniques have been maintained to be applied on tourism and hotel demand estimations and empiric evidences have displayed satisfying performances. Most of these studies are published in the magazines on other disciplines such as science and statistics calculation. A possible reason is that these studies have mostly focused on methodically development and assessment of estimation accuracy instead of tourism-specific practises. Moreover, establishing a model based on artificial intelligent is deprived of strong theoretical foundations and it is difficult to measure effect of economic factors on tourism and hotel demand by using such models. These explain why artificial intelligence based models have been limited to tourism and hotel demand analysis and number of publications on artificial intelligence is limited in tourism and accommodation

magazines. The artificial neutral network (ANN) is the artificial intelligence based technique most commonly seen in the recent literature. Other techniques such as support vector regression (SVR), rough set model, fuzzy system methods, genetic algorithms and Gauss process regression (GPR) have also been used for tourism and hotel demand estimated albeit in a less frequent way. (Wu etc. 2016).

Artificial Neural Network which is a nonparametric and data oriented technique has attracted great attention due to its ability of mapping the linear or nonlinear function without an assumption forced by the modelling process. Layers simulating the biological neural system, especially the human brain including input and output have one or more neurons.

These neurons are related to each other in information and data processing process. (Cuhadar etc. 2014) Different ANN models have been implemented for tourism and hotel estimation applications such as multilayer perceptron(MLP), radial basis function (RBF), general regression neural network (GRNN) and Elman neural network (Elman NN). MLP is the most commonly used ANN (Artificial Neural Network) model and has more three or more neural layers with non-linear activation function.(Chen etc. 2012), (Claveria etc. 2014), (Lin etc. 2011) SVR is another artificial intelligent based model. It minimizes the learning error by applying the structural risk principle in contrast with the ANN adopting the risk minimization principal. SVR solves the problems of linear regression by mapping input data on a large area non-linearly. Theoretically, SVR, a ANN model, could yield a global optimum level instead of limited optimisation.(Hong etc. 2011) SVR has been used in various studies on tourism and hotel estimations. (Cang 2014), (Chen etc. 2007), (Hong etc. 2011), (Xu etc. 2009)

Fuzzy system model is appropriate in cases in which data have been formed in terms of language or consists of less than 50 data points.(Tsaur etc. 2011) Different versions of the fuzzy stem model are used for tourism and hotel demand estimations. For example (Aladag etc. 2014), a fuzzy system model has been used to estimate the international tourism demand in Turkey by seasonal time series. (Chen etc. 2010)

It has been observed in the studies with SSCI and SCI indexes that SVR and MLP models have been comparatively studies and these models have been used in the recent studies on tourism accommodation estimations and modelling. However, it has been found out as a conclusion that there is not a method providing clear and precise results in these estimation studies. (Cankurt etc. 2016) Studies comparing Linear, MLP, and SVR models are available in the literature. Linear and non-linear models were compared in studies in the literature. Studies comparing regression analysis in tourism forecasts are available in the literature. (Cang 2014), (Cankurt etc. 2016)

2. DATA SET and MULTIVARIATE APPROACH for TOURISM DEMAND MODELLING

As used in many studies, annual time series analyses have been used to make future estimations by data set time series analyses and regression methods of past figures. The multiple variables related to estimation method, USD exchange rate, occupancy rates, arriving tourist numbers, average accommodation duration of tourists, number of overnight stays and CPI, have been used and implemented in the study by analysis via time series methods.

In the study, annual time series data of the variables have been cited from the websites of the following institutions: Republic of Turkey Ministry of Tourism(www.turizm.gov.tr), Turkish State Institute of Statistics(www.die.gov.tr), Central Bank of Republic of Turkey(http://evds.tcmb.gov.tr), European Central Bank(ECB)(https://www.ecb.europa.eu) and TURSAB(www.tursab.org.tr).

3. DATA and MODELLING APPROACHES

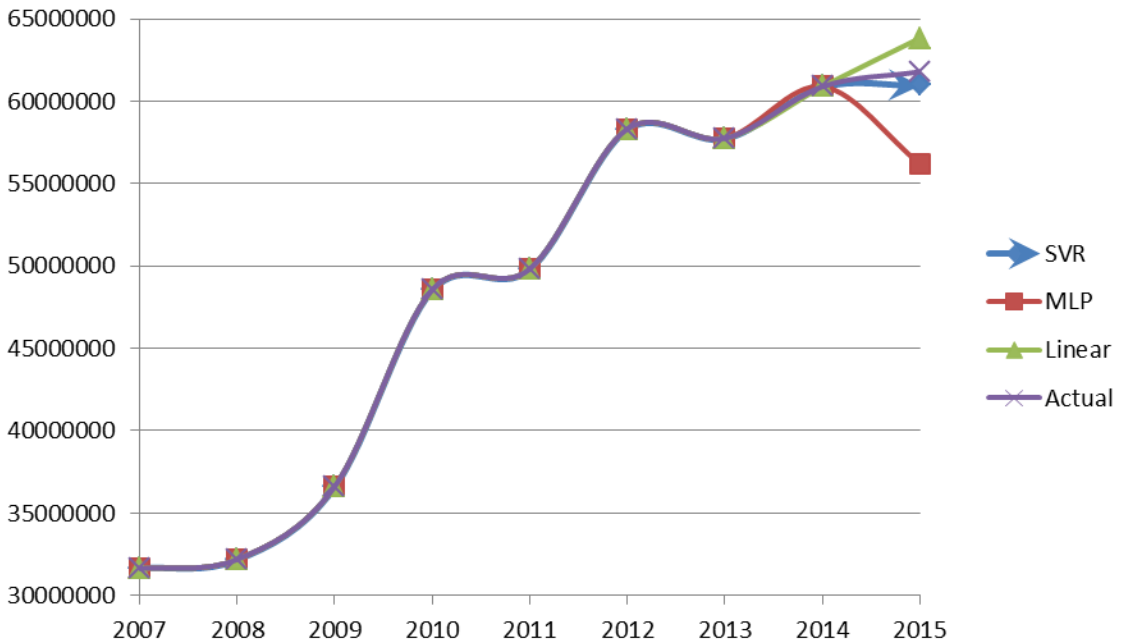
As seen from the studies performed on tourism demand estimation, the main factors affecting the tourism demand estimation reveal the relation between variables (USD exchange rate, occupancy rates, arriving tourist numbers, average accommodation duration of tourists, number of overnight stays, CPI) and demand and determinants explaining the data mining methods.

4. IMPLEMENTATION

Number of overnight stays in Antalya, Istanbul and Mugla, the most tourist-attracting cities of Turkey, has been estimated by adding domestic and foreign tourists and obtaining the grand total of the numbers arriving at these provinces. Then, estimation studies on number of domestic tourists have been performed. And estimation studies on number of domestic tourists have also been carried out. Numbers of two estimation studies have been added and they have been compared by using different regression methods. In the implementation stage, WEKA 3.8 data mining software has been used. Weka was originally developed by Waikato University in New Zealand. It has widespread use for machine learning and data mining. The state-of-the-art machine written in Java contains a large collection of algorithms for learning and data mining.WEKA includes regression, classification, clustering, association rules, visualization and data preprocessing tools. WEKA is very popular for academic and industrial researchers. And is widely used for teaching purposes. (Naik etc. 2016) WEKA is the most popular and open source data mining tool.(Duriqi etc. 2016) Correlation coefficient, Relative absolute error, root mean square error, Estimated Values and Real Values have been determined and results of regression analysis methods have been compared for the provinces of Antalya, Istanbul and Mugla, the most tourist-attracting cities of Turkey, in the study by

using Linear Regression, Multilayer Perception and Support Vector Regression analyses methods which are WEKA 3.8 data mining regression analyses methods and time series analyses. The tourism data from the years of 2016 and 2017 have not been used in this study due to terrorist attacks and crisis with other countries, Russia for particular. Data of 2007-2014 have been used. The forecast for the year of 2015 has been generated. And compared with the actual values of 2015.

Figure 1. Total Number of Overnight Stays in the Province of Antalya (Domestic and Foreign)



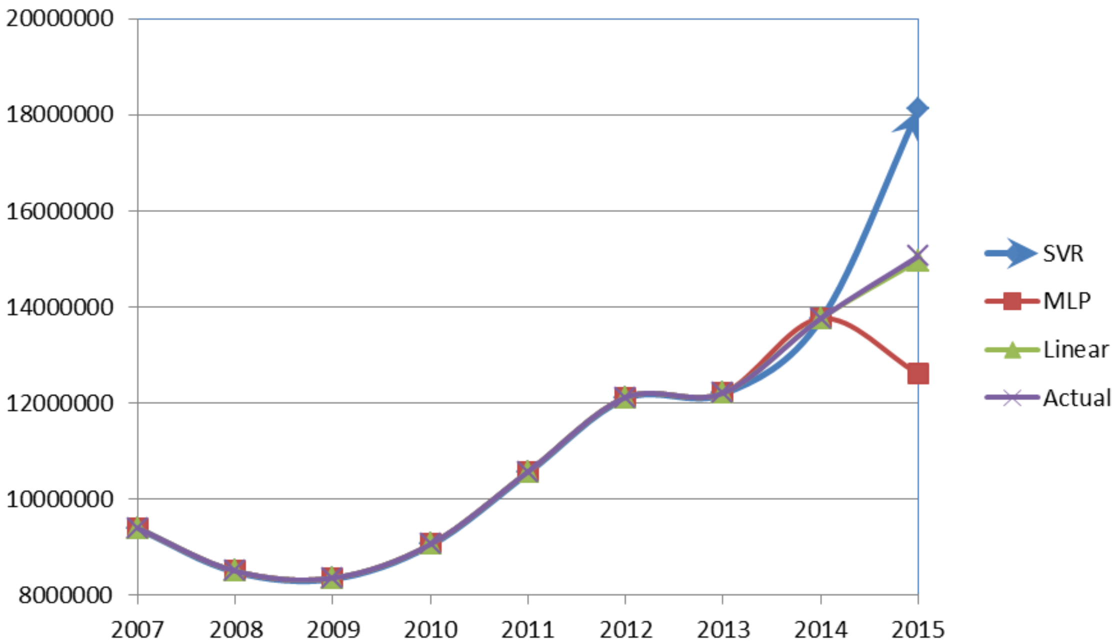
Comparison of all methods for the province of Antalya.

Table 1.Overall Performances of all the Methods for Antalya (Overnight Stays in 2015)

Model	Rank*	Correlation coefficient	Relative Absolute error	Root relative squared error	Forecasting Values	Actual Values
SVR	1	0.9888	8.8528%	15.1269%	60.985.186	61.799.160
MLP	3	0.9999	1.6364%	1.81%	56.195.674	61.799.160
Linear	2	0.9998	1.4558%	1.8981 %	63.840.965	61.799.160

*Ranking values have been determined by proximity of estimated values to actual values

Figure 2. Total Number of Overnight Stays in the Province of Istanbul(Domestic and Foreign)



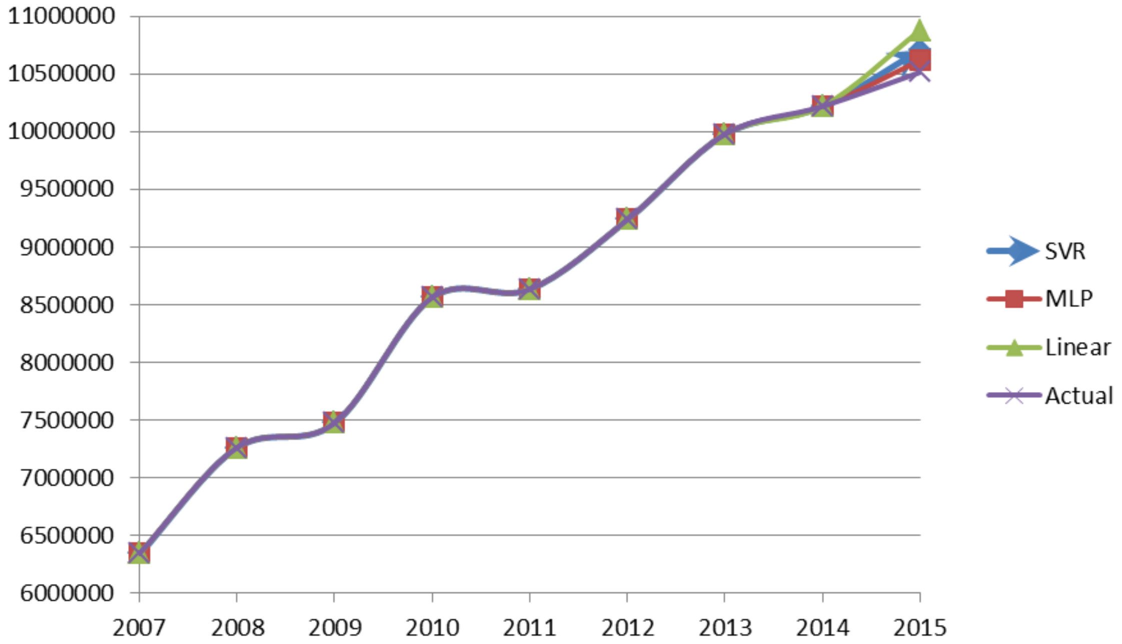
Comparison of all methods for the province of Istanbul

Table 2.Overall Performances of all the Methods for Istanbul (Overnight Stays in 2015)

Model	Rank*	Correlation coefficient	Relative Absolute error	Root relative squared error	Forecasting Values	Actual Values
SVR	3	0.9992	2.7128 %	4.6662%	18.119.499	15.054.304
MLP	2	1	0.0993%	0.1026%	12.610.537	15.054.304
Linear	1	0.9996	2.5562%	2.7829%	14.964.781	15.054.304

*Ranking values have been determined by proximity of estimated values to actual values

Figure 3. Total Number of Overnight Stays in the Province of Mugla (Domestic and Foreign)



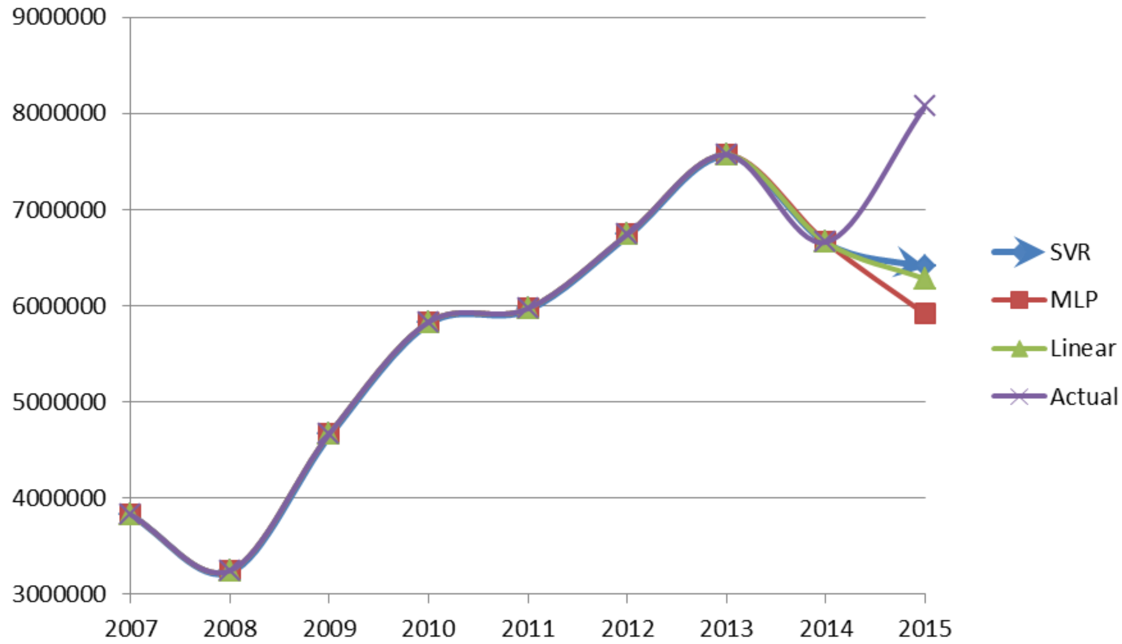
Comparison of all methods for the province of Mugla.

Table 3.Overall Performances of all the Methods for Mugla (Overnight Stays in 2015)

Model	Rank*	Correlation coefficient	Relative Absolute error	Root relative squared error	Forecasting Values	Actual Values
SVR	2	0.999	3.0693%	5.355%	10.708.283	10.515.144
MLP	1	1	0.0113%	0.0118%	10.618.341	10.515.144
Linear	3	0.9994	3.6654%	3.5552%	10.878.364	10.515.144

*Ranking values have been determined by proximity of estimated values to actual values

Figure 4. Estimation Models for Number of Overnight Stays of Domestic Tourists in the Province of Antalya



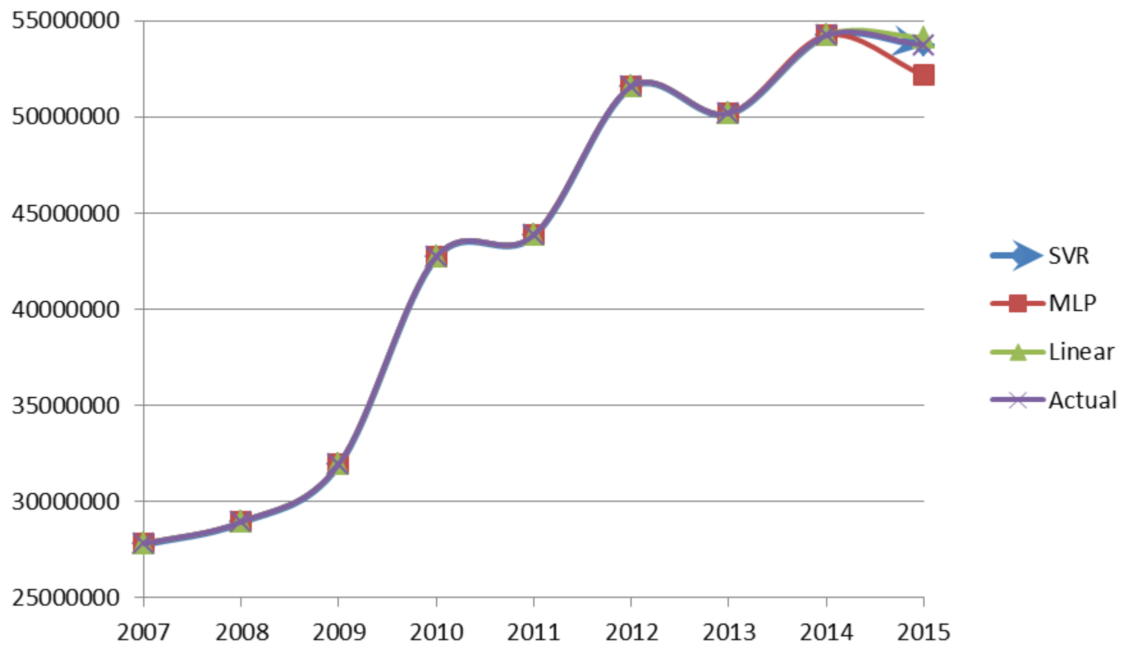
Comparison of all methods for the province of Antalya.(Domestic)

Table 4.Overall Performances of all the Methods for Antalya Domestic Tourists (Overnight Stays in 2015)

Model	Rank*	Correlation coefficient	Relative Absolute error	Root relative squared error	Forecasting Values	Actual Values
SVR	1	0.9968	4.8138 %	8.0087 %	6.411.406	8.070.524
MLP	3	1	0.0503 %	0.05 %	5.913.077	8.070.524
Linear	2	1	0.1502 %	0.1519 %	6.278.579	8.070.524

*Ranking values have been determined by proximity of estimated values to actual values

Figure 5. Estimation Models for Number of Overnight Stays of Foreign Tourists in the Province of Antalya



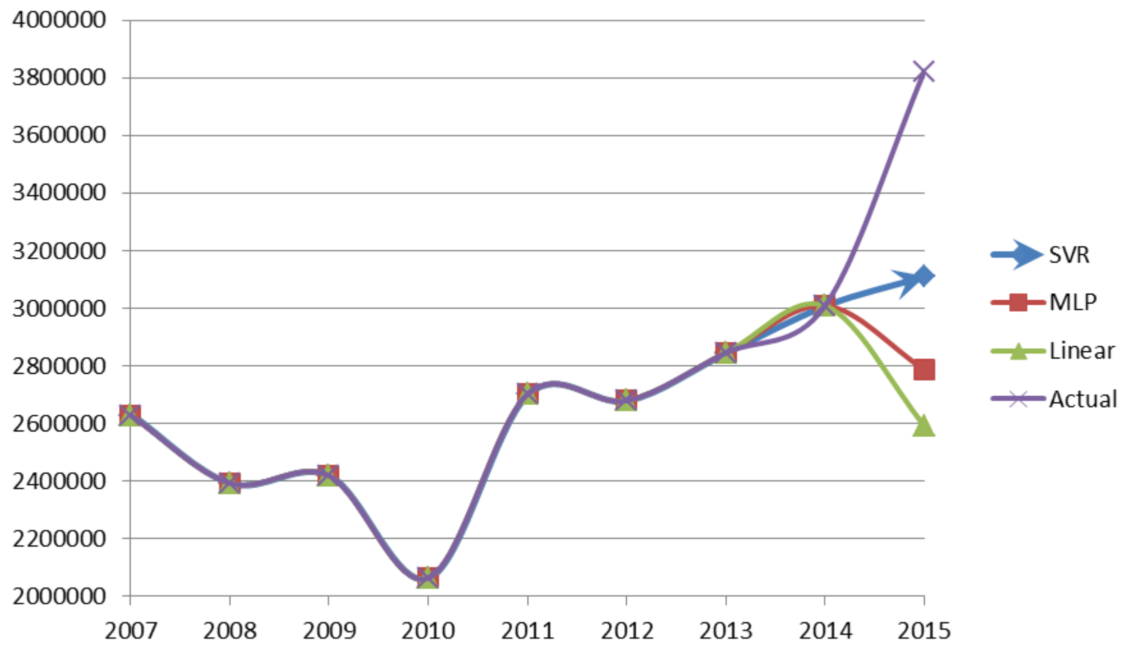
Comparison of all methods for the province of Antalya. (Foreign)

Table 5.Overall Performances of all the Methods for Antalya Foreign Tourists (Overnight Stays in 2015)

Model	Rank*	Correlation coefficient	Relative Absolute error	Root relative squared error	Forecasting Values	Actual Values
SVR	1	0.9879	9.7046 %	16.3781 %	53.735.407	53.728.636
MLP	3	1	0.6541 %	0.7441 %	52.153.886	53.728.636
Linear	2	0.9995	3.0687 %	3.0928 %	54.129.285	53.728.636

*Ranking values have been determined by proximity of estimated values to actual values

Figure 6. Estimation Models for Number of Overnight Stays of Domestic Tourists in the Province of Istanbul.



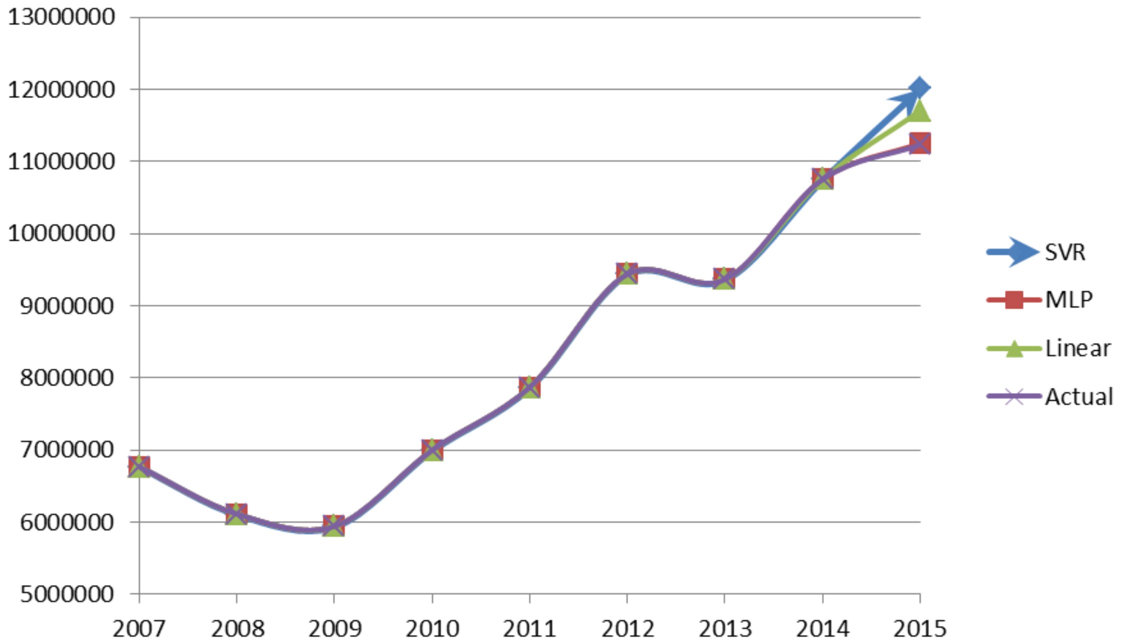
Comparison of all methods for the province of Istanbul(Domestic)

Table 6.Overall Performances of all the Methods for Istanbul Domestic Tourists (Overnight Stays in 2015)

Model	Rank*	Correlation coefficient	Relative Absolute error	Root relative squared error	Forecasting Values	Actual Values
SVR	1	0.9956	7.2806 %	10.6956 %	3.111.332	3.822.904
MLP	3	1	0.0337 %	0.033 %	2.786.599	3.822.904
Linear	2	0.9999	1.4895 %	1.5961 %	2.592.368	3.822.904

*Ranking values have been determined by proximity of estimated values to actual values

Figure 7. Estimation Models for Number of Overnight Stays of Foreign Tourists in the Province of Istanbul.



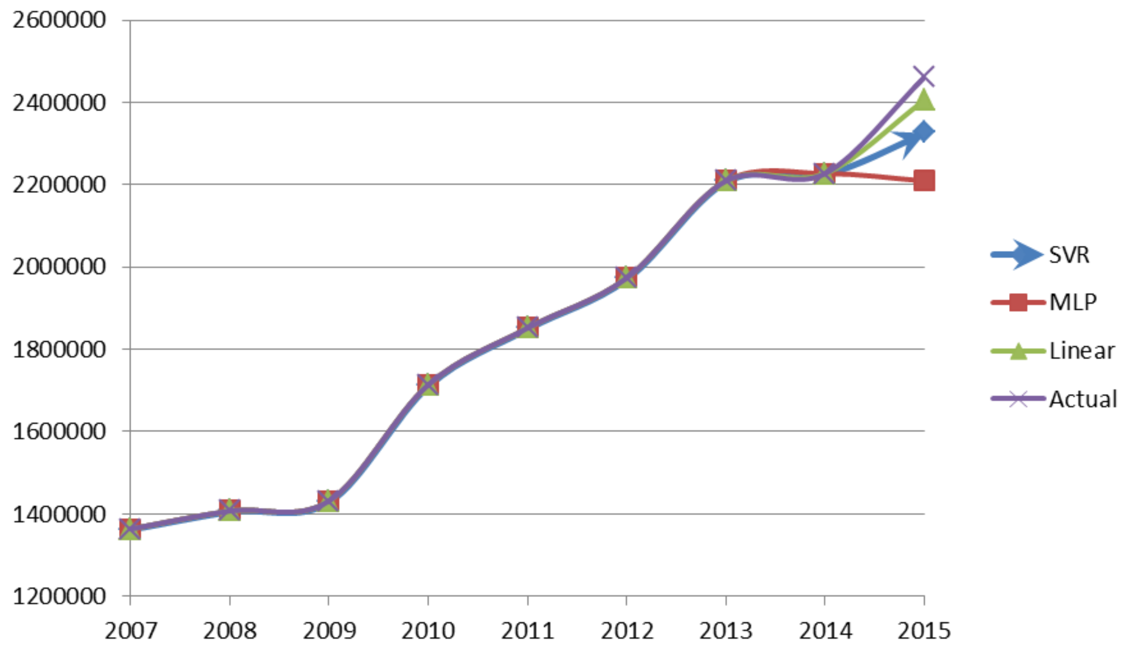
Comparison of all methods for the province of Istanbul (Foreign)

Table 7.Overall Performances of all the Methods for Istanbul Foreign Tourists (Overnight Stays in 2015)

Model	Rank*	Correlation coefficient	Relative Absolute error	Root relative squared error	Forecasting Values	Actual Values
SVR	3	0.998	3.1059 %	6.6263 %	12.011.135	11.231.400
MLP	1	1	0.7463 %	0.722 %	11.254.610	11.231.400
Linear	2	1	0.8295 %	0.8259 %	11.701.222	11.231.400

*Ranking values have been determined by proximity of estimated values to actual values

Figure 8. Estimation Models for Number of Overnight Stays of Domestic Tourists in the Province of Mugla.



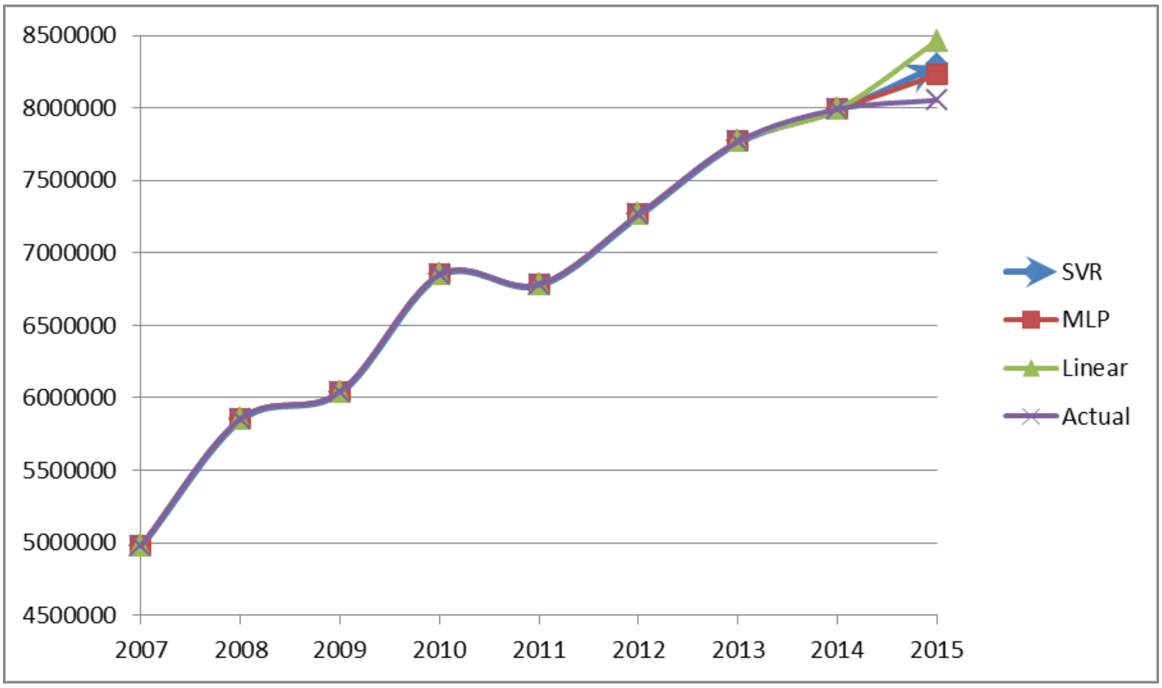
Comparison of all methods for the province of Mugla(Domestic)

Table 8.Overall Performances of all the Methods for Mugla Domestic Tourists (Overnight Stays in 2015)

Model	Rank*	Correlation coefficient	Relative Absolute error	Root relative squared error	Forecasting Values	Actual Values
SVR	2	0.9998	1.3415 %	2.3495 %	2.327.350	2.461.172
MLP	3	1	0.1078 %	0.1147 %	2.209.586	2.461.172
Linear	1	0.9998	2.0166 %	2.2093 %	2.405.603	2.461.172

*Ranking values have been determined by proximity of estimated values to actual values

Figure 9. Estimation Models for Number of Overnight Stays of Foreign Tourists in the Province of Mugla.



Comparison of all methods for the province of Mugla(Foreign)

Table 9.Overall Performances of all the Methods for Mugla Domestic Tourists (Overnight Stays in 2015)

Model	Rank*	Correlation coefficient	Relative Absolute error	Root relative squared error	Forecasting Values	Actual Values
SVR	2	0.9991	3.5406 %	4.3649 %	8.299.582	8.053.972
MLP	1	1	0.0259 %	0.0241 %	8.229.138	8.053.972
Linear	3	0.9994	3.3414 %	3.4669 %	8.463.889	8.053.972

*Ranking values have been determined by proximity of estimated values to actual values

5. CONCLUSIONS AND FUTURE CONSIDERATIONS

This paper investigates three forecasting methods; SVR, MLP and Linear. Three different destinations were studied; Antalya, Istanbul and Mugla.

In the graphs with high density of cyclical fluctuation, it was observed that SVR method gave the closest result to real values. In the graphs with less density from high of cyclical fluctuation, it was observed that MLP method gave the closest result to real values. In the graphs with low density of cyclical fluctuation, it was observed that Linear method gave the closest result to real values.

In future, more research will be focused to apply more sophisticated forecasting techniques using latest technology. Other advanced data mining techniques to predict visitor arrivals.

References:

1. Hemambika, P. (2005). "Data Mining Framework Dissertation"
2. Kadambi, R.R.(2005) "Analysis of Data Mining Techniques for Customer Segmentation and Predictive Modeling-A Case Study"
3. Lim, J. (2013). "An efficient approach to clustering data sets with mixed type attributes in datamining."
4. Cloyd,J. D. (2002)."Data Mining with Newton's Method" (Doctoral dissertation, East Tennessee State University).
5. Al-Alwani, M. A. (2014). "Data mining and statistical analysis of completions in the Canadian Montney formation."
6. Ibrahim, A. M. (1995). "Forecasting Hotel Occupancy Rates in Egypt Using Time Series Models."
7. Rakthanmanon, T.(2012). "Efficient Algorithms for High Dimensional Data Mining"
8. Seliem,M. L. (2006). "Foreign Exchange Forecasting Using Artificial Neural Network as Data Mining Tool."
9. Yunyue, Z. (2004). "High Performance Data Mining in Time Series: Techniques and Case Studies" (Doctoral dissertation, PHD. Thesis). New York City: NY University, 2004.138 139).
10. Siripitayananon, P. (2002). "Data mining techniques for handling a missing data problem." The University of Alabama.
11. Ding, Y. (2014). "Estimating truncated hotel demand: A comparison of low computational cost forecasting methods" (Doctoral dissertation, Purdue University).
12. Gaojun, L., and Boxue, L. (2009, December). "The research on Combination forecasting model of the automobile sales forecasting system." In Computer Science-Technology and Applications, 2009. IFCSTA'09. International Forum on (Vol. 3, pp.

82-85). IEEE.
13. Yu, X.,Qi, Z., and Zhao, Y. (2013). "Support Vector Regression for Newspaper/Magazine Sales Forecasting." Procedia Computer Science, 17, 1055-1062.
14. Kong, X. etc. (2015). "Wind speed prediction using reduced support vector machines with feature selection." Neurocomputing, 169, 449-456.
15. Dai, W.,Chuang, Y. Y., and Lu, C. J. (2015). "A Clustering-based Sales Forecasting Scheme Using Support Vector Regression for Computer Server." Procedia Manufacturing, 2, 82-86.
16. García, F. T.,Villalba, L. J. G., and Portela, J. (2012). "Intelligent system for time series classification using support vector machine sapplied to supply-chain." Expert Systems with Applications, 39(12), 10590-10599.
17. Chern, C. etc. (2015). "A sales forecasting model for consumer products based on thein fluence of online word-of-mouth." Information Systems and e-Business Management, 13(3), 445-473.
18. Fang, D., and Weng, W. (2011). "Sales Forecasting System for Chinese Tobacco Wholesalers." Procedia Environmental Sciences, 11, 380-386.
19. Moon, S.,Bae, S., and Kim, S. (2014, August). "Predicting the Near-Weekend Ticket Sales Using Web-Based External Factorsand Box-Office Data." In Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 02(pp. 312-318). IEEE Computer Society.
20. Tezel, G., and Buyukyildiz, M. (2016). "Monthly evaporation forecasting using artificial neural networks and support vector machines." The oretical and Applied Climatology, 124(1-2), 69-80.
21. DiPillo, G. etc. (2016). "An application of support vector machines to sales forecasting under promotions." 4OR, 1-17.
22. Dai, W.,Wu, J. Y., and Lu, C. J. (2014). "Applying different independent component analysis algorithms and support vector regression for IT chain store sales forecasting." The Scientific World Journal, 2014.
23. Lu, C. J. (2014). "Sales forecasting of computer products based on variable selections cheme and support vector regression." Neurocomputing, 128, 491-499.
24. Zhang, F. etc. (2016). "Time series forecasting for building energy consumption using weighted Support Vector Regression with different ialevolution optimization technique." Energy and Building, 2016
25. Zhao, W. Tao, T. and Zio, E.(2013). "Parameters Tuning in Support Vector Regression for Reliability Forecasting." Chemical Engineering Transactions, 2013
26. Gomez, D. etc.(2011)."Improving sale performance prediction using support vector machines." Expert Systems with Applications, 2011

27. Khalil Zadeh, N.,Sepehri, M. M., and Farvaresh, H. (2014). "Intelligent Sales Prediction for Pharmaceutical Distribution Companies: A Data Mining Based Approach." Mathematical Problems in Engineering, 2014.
28. Cankurt, S., and Subasi, A. (2016) "Tourism demand modelling and forecasting using data mining techniques in multivariate time series: a casestudy in Turkey."
29. Wu, D.,Song, H., and Shen, S. (2016). "New Developments in Tourism and Hotel Demand Modeling and Forecasting." International Journal of Contemporary Hospitality Management.
30. Koupriouchina, L.,van der Rest, J. P., and Schwartz, Z. (2014). "On revenue management and theuse of occupancy forecasting error measures." International Journal of Hospitality Management, 41, 104-114.
31. Lim, C.,Chang, C., and McAleer, M. (2009). "Forecasting h (m) otel guest nights in New Zealand." International Journal of Hospitality Management, 28(2), 228-235.
32. Cortés-Jiménez, I., and Blake, A. (2011). "Tourism demand modeling by purpose of visit and nationality." Journal of Travel Research, 50(4), 408-416.
33. Smeral, E. (2010). "Impacts of the world recession and economic crisis on tourism: Forecasts and potential risks." Journal of Travel Research, 49(1), 31-38.
34. Athanasopoulos, G., and Hyndman, R. J. (2008). "Modelling and forecasting Australian domestic tourism." Tourism Management, 29(1), 19-31.
35. Baggio, R., and Sainaghi, R. (2016). "Mapping time series in to networks as a tool to assess the complex dynamics of tourism systems." Tourism Management, 54, 23-33.
36. Cuhadar, M.,Cogurcu, I., and Kukrer, C. (2014). "Modelling and forecasting cruise tourism demand to Izmir by different artificial neural network architectures." International Journal of Business and Social Research, 4(3), 12-28.
37. Guizzardi, A.,and Stacchini, A. (2015). "Real-time forecasting regional tourism with business sentiment surveys." Tourism Management, 47, 213-223.
38. Falk, M. (2014). "Impact of weather conditions on tourism demand in the peak summer season over the last 50 years." Tourism Management Perspectives, 9, 24-35.
39. Corgel, J.,Lane, J., and Walls, A. (2013). "How currency exchange rates affect the demand for US hotel rooms." International Journal of Hospitality Management, 35, 78-88.
40. Song, H. etc. (2011). "Impact of financial/economic crisis on demand for hotel rooms in Hong Kong." Tourism Management, 32(1), 172-186.
41. Wu, E. H.,Law, R., and Jiang, B. (2010). "Data mining for hotel occupancy rate: an independent component analysis approach." Journal of Travel &Tourism Marketing, 27(4), 426-438.
42. Chen, M. H. (2013). "Determinants of the Taiwanese tourist hotel industry cycle." Tourism Management, 38, 15-19.
43. Zheng, T. (2014). "What caused the decrease in Rev PAR during the recession? An

ARIMA with intervention analysis of room supply and market demand." International Journal of Contemporary Hospitality Management, 26(8), 1225-1242.
44. Croes, R., and Semrad, K. J. (2012). "Discounting works in the hotel industry: a structural approach to understanding why." Tourism Economics, 18(4), 769-779.
45. Chen, C. F.,Lai, M. C., and Yeh, C. C. (2012). "Forecasting tourism demand based on empirical mode decomposition and neural network." Knowledge-BasedSystems, 26, 281-287.
46. Claveria, O., and Torra, S. (2014). "Forecasting tourism demand to Catalonia: Neural networks vs. time series models." Economic Modelling, 36, 220-228.
47. Lin, C. J.,Chen, H. F., and Lee, T. S. (2011). "Forecasting tourism demand using time series, artificial neural networks and multivariate adaptive regression splines: Evidence from Taiwan." International Journal of Business Administration, 2(2), 14.
48. Hong, W. C., etc. (2011). "SVR with hybrid chaotic genetic algorithms for tourism demand forecasting." AppliedSoft Computing, 11(2), 1881-1890.
49. Tsaur, R. C., and Kuo, T. C. (2011). "The adaptive fuzzy time series model with an application to Taiwan's tourism demand." Expert systems with Applications, 38(8), 9164-9171.
50. Cang, S. (2014). "A Comparative Analysis of Three Types of Tourism Demand Forecasting Models: Individual, Linear Combination and Non-linear Combination." International Journal of Tourism Research, 16(6), 596-607.
51. Aladag, C. H., etc. (2014). "A high order seasonal fuzzy time series model and application to international tourism demand of Turkey." Journal of Intelligent&FuzzySystems, 26(1), 295-302.
52. Chen, K. Y., and Wang, C. H. (2007). "Support vector regression with genetic algorithms in forecasting tourism demand." Tourism Management, 28(1), 215-226.
53. Chen, M. S.,Ying, L. C., and Pan, M. C. (2010). "Forecasting tourist arrivals by using the adaptive network-based fuzzy inference system." Expert Systemswith Applications, 37(2), 1185-1191.
54. Xu, X.,Law, R., and Wu, T. (2009). "Support vector machines with manifold learning and probabilistic space projection for tourist expenditure analysis." International Journal of Computational Intelligence Systems, 2(1), 17-26.
55. Republic of Turkey Ministry of Tourism. www.turizm.gov.tr (Accessed: 15.11.2017)
56. Turkish State Institute of Statistics. www.die.gov.tr (Accessed: 30.11.2017)
57. Central Bank of Republic of Turkey. http://evds.tcmb.gov.tr (Accessed: 31.11.2017)
58. European Central Bank (ECB). https://www.ecb.europa.eu (Accessed: 10.12.2017)
59. Association of Turkish Travel Agencies(TURSAB)www.tursab.org.tr (Accessed: 30.11.2017)
60. Naik, A., and Samant, L. (2016). "Correlation review of classification algorithm using data mining tool: WEKA, Rapidminer, Tanagra, Orange and Knime." Procedia

Computer Science, 85, 662-668.

61. Duriqi, R., Raca, V., and Cico, B. (2016, June). "Comparative analysis of classification algorithms on three different datasets using WEKA." In Embedded Computing (MECO), 2016 5th Mediterranean Conference on (pp. 335-338). IEEE.